Description

TWIN EEPROM MEMORY TRANSISTORS WITH
SUBSURFACE STEPPED FLOATING GATES

5

CROSS-REFERENCE TO RELATED APPLICATION

This application is a continuation-in-part of prior application Serial No. 10/423,637 filed April 25, 2003, a continuation-in-part of prior application Serial

10   No. 10/465,718 filed June 18, 2003, and a continuation-in-part of prior application Ser. No. 10/680,355, filed October 6, 2003. All three applications are herewith incorporated by reference in their entirety.

15   ## TECHNICAL FIELD

The invention relates to non-volatile memory transistors and, in particular, to a compact arrangement of such memory cells for an array and a method of making them.

20

BACKGROUND ART

In prior application Serial No. 10/423,637 entitled "Mirror Image Memory Cell Transistor Pairs Featuring Poly Floating Spacers," as well as in prior

25   application Serial No. 10/465,718 entitled "Mirror Image Non-Volatile Memory Cell Transistor Pairs with Single Poly Layer," both assigned to the assignee of the present invention, B. Lojek described an arrangement of non-volatile MOS memory transistors for a memory array

30   wherein symmetric pairs of transistors were built in a memory array. Transistor pairs shared an electrode in a common well, such as a drain electrode, but were otherwise completely independent. The pair was manufactured between a pair of isolation regions and

35   sharing the same substrate region, almost as if a single transistor were constructed there.

In the prior art, single MOS floating gate
transistors that stored two data bits have been devised
as a way to achieve compactness. Since millions of data
bits are frequently stored in non-volatile memory arrays,

5   small savings of space are multiplied significantly over
the array. In prior application Ser. No. 10/327,336
entitled "Multi-Level Memory Cell with Lateral Floating
Spacers," assigned to the assignee of the present
invention, B. Lojek described how two spacers, on

10   opposite sides of a conductive gate, behave as
independent charge storage regions for separate binary
data, thereby allowing a single non-volatile MOS
transistor to store two binary bits. Each memory cell is
connected to two bit lines and one word line. The bit

15   lines are phased so that during a single clock cycle,
first one bit line is active and then the other while a
word line is active for the entire cycle. In this
manner, both storage areas may be accessed for a read or
write operation in a single clock cycle.

20   In U.S. Patent No. 6,043,530 to M. Chang, a MOS
memory transistor construction is shown employing band-
to-band tunneling. In U.S. Patent No. 6,323,088 to F.
Gonzalez et al., a multibit charge storage transistor
addressing scheme is shown with phased bit lines.

25   In the prior art, multibit charge storage
structures are known that achieve good data density in a
memory array without giving up valuable chip space. One
of the problems that is encountered as density increases
is that the amount of crosstalk between storage sites

30   increases. Because the charge storage structures are so
small, one charge storage location can sometimes
influence another. On the other hand, separation of
charge storage sites gives up chip space. The ultimate
separation is one dedicated transistor for each data bit.

35   Accordingly, an object of the invention is to provide
good separation for data bits afforded by dedicated

transistors yet achieve the compactness of multibit charge storage structures for a non-volatile memory array.

5   SUMMARY OF THE INVENTION

The above object has been achieved with a memory array having cells with twin EEPROM memory transistors that occupy a space almost the same size as a single EEPROM memory transistor. The twin transistors of
10   each cell are symmetrically arranged in a common substrate and feature a single poly layer, with portions used as floating gates that are stepped below the level of the substrate surface, yet insulated from the substrate by thin oxide. The floating gate electrically
15   communicates with a subsurface electrode that participates in charge transfer to the floating gate. The usual EEPROM control gate is replaced by a first capacitor wherein the same poly portion used to form the floating gate extends to form a second plate of the first
20   capacitor. The first plate of the first capacitor is a control line connected to a phased signal source whereby phasing of plates of the twin cells allows each transistor to act independently. The drain of each transistor is connected to one plate of a second
25   capacitor and to a bit line while the second plate is connected to a word line.

By stepping the floating gate into the substrate and forming a floating gate corner in the substrate, the intensity of electric field from a
30   subsurface electrode will increase and enhance tunneling action.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is an electrical schematic drawing of
35   memory cells forming the core of a memory array of the present invention.

Fig. 2 is a side sectional view of an early
manufacturing step for a memory transistor in the memory
cells of Fig. 1.

Fig. 3 is a top view of a mask for making a
5   substrate step shown in the sectional view of Fig. 2.

Fig. 4 is a top view of a layout in an early
manufacturing stage of twin memory cells shown in Fig. 1.

Fig. 5 is a side sectional view taken along
lines 5-5 in Fig. 4.

10   Fig. 6 is a side sectional view taken along
lines 6-6 in Fig. 4.

Fig. 7 is a top view of a layout in an
intermediate manufacturing stage of twin memory cells
shown in Fig. 1.

15   Fig. 8 is a side sectional view taken along
lines 8-8 in Fig. 7.

Fig. 9 is a side sectional view following Fig.
8 at a later stage in manufacturing.

Fig. 10 is a top view of a contact mask
20   superposed on the top view of Fig. 7, with conductor
shading for the single polysilicon layer.

Fig. 11 is an electrical schematic drawing of
twin symmetric memory cells shown in Fig. 1 redrawn for
comparison with Fig. 10, including locations of contacts
25   shown in Fig. 10.

BEST MODE FOR CARRYING OUT THE INVENTION

With reference to Fig. 1, a memory cell 13 in a
memory array 10 is seen to have first and second non-
30   volatile memory transistors 15 and 115, respectively.
The first memory transistor 15 has a drain 21 connected
to select capacitor 19, a floating gate 23 connected to
control capacitor 29 and a source 25 connected to the
source contact 27.

35   Select capacitor 19 has a first electrode 31
connected to drain 21 of first memory transistor 15 and

also connected to the first bit line, BL1. The second electrode 33 of select capacitor 19 is connected to word line WL1. The word line WL1 is extended from first electrode 31 along line 35 into another cell. The

5   floating gate 23 of memory transistor 15 is connected to a first electrode 37 of control capacitor 29, while second electrode 39 is connected to a first control line terminal 41. A pulse on terminal 41 charges the second electrode 39, causing induced charge to appear on first

10   electrode 37 which forms a floating gate together with electrode 23. This is one of two ways in which charge appears on the floating gate 23. Another way for charge to appear is by tunneling or electron injection from source or drain electrode 21 and 25. When one voltage is

15   applied to bit line BL1 and another voltage is applied at source contact 27 charge may be transferred onto the floating gate 23 by tunneling charge transfer mechanisms. Just as the word line WL1 extends into another memory cell in the same column along line 35, bit line BL1 is

20   also extended into a memory cell in the same row along line 43.

       The second memory transistor 115 is symmetric with first memory transistor 115 relative to source contact 27. The second memory transistor 115 has a

25   floating gate 123 which may be charged by control capacitor 129. Memory transistor 115 has a drain electrode 121 connected to a first plate 131 of select capacitor 119 and a source electrode 125 connected to source contact 27. The first plate 131 is also connected

30   to the bit line BL1. The second plate of capacitor 119 is connected to the word line WL2. The word line WL2 is extended from the first electrode 131 along line 135 to a control capacitor (not shown) into a neighboring cell in the same column. The bit line 43 similarly extends from

35   the first electrode of select capacitor 119 into a neighboring cell in the same row.

Memory cell 13 is typical of the memory cells in the memory array 10. Each cell is seen to have twin non-volatile memory transistors that are symmetric about a source contact, such as source contact 27. The two memory transistors have floating gates associated with two control capacitors on the one hand and have drain or source electrodes associated with two select capacitors on the other hand. The two control line terminals 41 and 141 associated with the control capacitors allow programming of the two memory transistors so that each transistor is independent of the other, even though they share a common source electrode at source contact 27 and also share bit line BL1. Memory cell 13 is associated with two word lines, WL1 and WL2, as well as one bit line, BL1.

With reference to Fig. 2, a silicon p-type wafer provides a substrate doped to have a p-well with a surface 56, upon which a thin layer of oxide 57 is grown. The oxide layer has a thickness of approximately 100 angstroms. The oxide is covered with a thick photoresist layer 51 and then patterned with a mask 52, shown in Fig. 3. The mask is approximately square with a dimension near the lower limit of resolution of photolithography. The photoresist is then etched so that well-defined steps 53 and 54 form a depression 58 with upper and lower corners to a depth of approximately 500 angstroms below the substrate surface 56. The facing corners of steps 53 and 54 will enhance the electric field near the floating gates of twin memory transistors extending into the planar surface of the wafer. Corners at the top and bottom of each step are important for increasing electric field intensity to enhance tunneling. The floating gates are built upon the steps but insulated from the substrate by gate oxide 57.

With reference to Fig. 4, a mask set is shown defining the active regions of two memory cells. The

mask set includes masks 52 and 55 for defining common source electrodes of twin EEPROMs and masks 62 and 64, as well as masks 66 and 68 for defining control lines. Two linear masks define parallel bit lines BL1 and BL2. The

5 areas surrounding the masks are isolated by a shallow trench isolation, as shown in fig 5. Trenches in p-well or p-substrate 50 (Fig. 6) of a p-type silicon wafer substrate are filled with dielectric insulator material 72, 74, 76, 78, and 80 (Fig. 5), typically silicon

10 dioxide. The areas that are not dielectric material are subject to doping either by diffusion or implantation. This allows the memory cells to have diffused bit lines BL1 and BL2.

Turning to Fig. 6, when doping of subsurface

15 regions is complete, the substrate is coated with oxide, previously described in Fig. 2 but not shown in Fig. 6, and the depression 58 is formed below the surface 56 of substrate p-well substrate 50. The depression 58 has steps or corners 53 and 54 that will form part of

20 floating gates of memory transistors. The steps or corners 53 and 54 may be seen in Fig. 4 also.

With reference to Fig. 7, the diffused regions previously described with reference to Fig. 4 may be seen. The diffused regions include the areas where

25 source masks 52 and 55 as well as the control line diffusions 62, 64, 66, and 68. The diffused bit lines BL1 and BL2 are also seen. All of these structures lie below the surface of the p-well, or p-substrate, including steps or corners 53 and 54.

30 A layer of poly is deposited over the substrate surface and then etched leaving floating members 82, 84, 86, and 88. Portions of these floating members will become floating gates of twin EEPROM transistors. The floating members have portions extending over the control

35 line diffusions 62 and 64, as well as control line diffusions 66 and 68. Portions of the floating members

also extend over the source mask regions 52 and 55. The poly layer is also used to define word lines WL1 and WL2, spaced apart and lying outwardly of the cell core.

In Fig. 8, the p-well substrate 50 is seen with gate oxide layer 57 over the substrate surface including the depression 58. The poly layer deposited over the substrate has portions which define floating gates 82 and 84 that follow the contour of steps or corners 53 and 54. Outwardly of the floating gate regions 82 and 84 are poly word lines WL1 and WL2.

Fig. 9 follows Fig. 8 at a further point in the manufacturing process. Subsurface implants have been made in p-well substrate 50, particularly source implant 92, as well as drain implants 94 and 96. The subsurface bit line diffusions BL1 are also seen. The poly floating gates 82 and 84 have sidewall spacers, such as sidewall spacers 83 and 85 surrounding floating gate 82. Similarly, word lines WL1 and WL2 have sidewall spacers such as spacers 87 and 89 associated with word line WL1. After the spacers have been constructed, a layer of interlayer dielectric, ILD layer 101, is deposited over the poly one layer. The ILD layer 101 is masked and etched to create holes that allow insertion of metal contacts 102, 104, and 106. These conductive metal contacts make contact with subsurface regions. Metal contacts 102 and 106 contact the diffused bit line BL1. Metal contact 104 contacts a common subsurface electrode 92. The relation of the metal contacts with the top view of Fig. 7 may be seen in Fig. 10.

In Fig 10, the position of metal contacts 102, 104, and 106 may be seen. Also, contacts 112 and 114, associated with the control line diffusions 62 and 64, may be seen. Contact 104 is located in the center of mask 52 that defines a common electrode for twin side-by-side memory transistors. In other words, contact 104 is located at a plane of symmetry for the twin EEPROM

transistors. In Fig. 10, the single poly layer has been shaded, with portions of the layer forming poly members 82 and 84, defining the contoured floating gates associated with the subsurface steps toward the common

5   source electrode. Other portions of the poly one layer define the word lines WL1 and WL2, as indicated by shading. Note that the poly members 82 and 84 extend over the control line diffusions 62 and 64. These control line diffusions have metal contacts 112 and 114,

10   respectively.

In Fig. 11 positions of the contacts of the memory cell in the top view of Fig. 10 are indicated relative to an electrical schematic of a memory cell as shown in Fig. 1. A total of five contacts is used for

15   each cell with two contacts, 102 and 106, being on the bit line BL1. The contact 104 is associated with the common source between the twin symmetric memory transistors. The contacts 112 and 114 are associated with capacitors 29 and 129. Fig. 11 may be projected

20   upwardly, towards Fig. 10, where a rough comparison can be made of the various circuit elements. In Fig. 10, the word line WL1 is seen to overlie the bit line BL1 but spaced apart by insulator thereby forming capacitor 19 in Fig. 11. Similarly, a portion of poly member 82 is seen

25   to overlie control line diffusion 62 thereby forming capacitor 29 in Fig. 11.